SUPPLEMENTARY MATERIAL

Probabilistic Sensitivity Analysis of Biochemical Reaction Systems

Hong-Xuan Zhang,[†] William P. Dempsey, Jr.,[‡] and John Goutsias[†]

The Whitaker Biomedical Engineering Institute, The Johns Hopkins University, Baltimore, MD 21218, and Division of Bioengineering, California Institute of Technology, Pasadena, CA 91125

[†]The Whitaker Biomedical Engineering Institute, The Johns Hopkins University, Baltimore, MD 21218
[‡]Division of Bioengineering, California Institute of Technology, Pasadena, CA 91125

A. System Response Characteristics

It has been demonstrated in the literature that differences in the duration and strength of the activity of the doubly phosphorylated extracellular signal-regulated kinase (ERK-PP), produced by the mitogen activated protein kinase (MAPK) cascade considered in Section VI of the Main Text, may generate distinct biological outcomes, such as cell differentiation, proliferation, and apoptosis.[1–5] Experimental evidence suggests that immediate early gene (IEG) products function as sensors for ERK-PP signal duration and strength.[2,3] Moreover, it has been experimentally demonstrated that the time-integrated ERK-PP response directly correlates with DNA synthesis.[6,7]

A conceptually simple way to explain these results is to assume the presence of a relatively unstable IEG product C, which is phosphorylated by ERK-PP resulting in stable molecules C-P. If we use the reaction

$$\text{C} + \text{ERK-PP} \xrightarrow{k} \text{C-P} + \text{ERK-PP}$$

to model phosphorylation of C, then the concentration profile $x_{\text{C-P}}(t)$ of C-P will satisfy the following differential equation:

$$\frac{dx_{\text{C-P}}(t)}{dt} = kx_{\text{C}}(t)x_{\text{ERK-PP}}(t).$$

As a consequence,

$$x_{\text{C-P}}(t) = kc\int_0^t x_{\text{ERK-PP}}(\tau)\,d\tau, \qquad \text{for } t \geq 0,$$

where we assume for simplicity that the concentration of C remains constant for every $t \geq 0$. This shows that, at time t, the concentration of the stable phosphorylated product C-P will be proportional to the cumulative concentration of ERK-PP within the time interval $[0, t]$, in agreement with previously published results.[7]

The concentration of C-P can induce distinct biological outcomes by influencing transcriptional control. As a consequence, the integrated ERK-PP response is an important signalling characteristic for sensitivity analysis. It has been observed however that, due to certain biochemical factors (such as degradation and nuclear translocation), the integrated response of C-P may not be the only

factor affecting cellular response. As a matter of fact, experimental evidence suggests that different biological outcomes may be produced by an activated MAP kinase, such as ERK-PP, depending on whether its concentration remains above a critical level for a sufficient period of time.[1,6,8] Therefore, the duration and strength of ERK-PP activity are two additional signaling characteristics of importance to sensitivity analysis. Note however that Eq. (4) in the Main Text implies

$$I := \int_0^t x_{\text{ERK-PP}}(\tau)\, d\tau = t_0 \left[\frac{1}{t_0} \int_0^{t_0} x_{\text{ERK-PP}}(\tau)\, d\tau \right] = D \times S, \qquad \text{for } t \geq t_0,$$

which shows that the integrated response I, duration D, and strength S of ERK-PP depend on each other. To avoid redundancy, we choose in this paper the duration and strength as two independent signaling characteristics of interest to sensitivity analysis.

If T is the timing of the ERK-PP profile, then $\int_0^T x_{\text{ERK-PP}}(\tau)\, d\tau = \int_T^{t_0} x_{\text{ERK-PP}}(\tau)\, d\tau$ [recall Eq. (4) in the Main Text]. As a result,

$$x_{\text{C-P}}(T) = \frac{x_{\text{C-P}}^{\max}}{2},$$

where

$$x_{\text{C-P}}^{\max} = \kappa c \int_0^{t_0} x_{\text{ERK-PP}}(\tau)\, d\tau$$

is the maximum attainable concentration of C-P. Therefore, T is the time required for the concentration of C-P to reach its half maximum value. Smaller values of T indicate a fast response of C-P activity to input stimuli, which may result in rapid activation of downstream transcriptional events that can significantly influence cellular response. We therefore believe that the timing, as defined by Eq. (4) in the Main Text, may be another important characteristic of the ERK-PP activity profile that is relevant to sensitivity analysis.

B. Derivative Approximation of Variances

If we assume that the response function $R(\boldsymbol{u})$ is sufficiently smooth around $\mathbf{0}$, so that its derivatives of order ≥ 3 at $\mathbf{0}$ are negligible, then its Taylor series expansion about $\mathbf{0}$ is given by

$$R(\boldsymbol{U}) \simeq R(\mathbf{0}) + \sum_{j=1}^{J} \frac{\partial R(\mathbf{0})}{\partial u_j} U_j + \frac{1}{2} \sum_{j=1}^{J} \sum_{j'=1}^{J} \frac{\partial^2 R(\mathbf{0})}{\partial u_j \partial u_{j'}} U_j U_{j'}. \tag{1}$$

From Eq. (1) above, we obtain

$$\mathrm{E}[R(\boldsymbol{U})] \simeq R(\mathbf{0}), \tag{2}$$

$$V_j = \mathrm{Var}[\mathrm{E}[R(\boldsymbol{U}) \mid U_j]] \simeq \lambda_j^2 \left[\frac{\partial R(\mathbf{0})}{\partial u_j} \right]^2, \tag{3}$$

and

$$V = \mathrm{Var}[R(\boldsymbol{U})] \simeq \sum_{j=1}^{J} \lambda_j^2 \left[\frac{\partial R(\mathbf{0})}{\partial u_j} \right]^2. \tag{4}$$

To show Eqs. (2)–(4) above, we use Eq. (26) in the Main Text and the fact that the factors U_j are statistically independent zero-mean Gaussian random variables with standard deviations λ_j. Recall that, for a zero-mean Gaussian random variable U_j with standard deviation λ_j, we have that

$$\mathrm{E}[U_j^3] = 0 \qquad \text{and} \qquad \mathrm{E}[U_j^4] = 3\lambda_j^4.$$

Eq. (27) and Eq. (28) in the Main Text are now a consequence of Eq. (3) and Eq. (4) above.

Note that, if $R(\boldsymbol{u})$ is linear, then Eqs. (2)–(4) above, and therefore Eq. (27) and Eq. (28) in the Main Text, will be satisfied with equality, regardless of the statistical model assumed for the factors U_j. More generally, Eq. (27) in the Main Text is satisfied with equality when $R(\boldsymbol{u})$ is additive. Indeed, if

$$R(\boldsymbol{u}) = \sum_{j=1}^{J} R_j(u_j),$$

then

$$V_j = \mathrm{Var}[\mathrm{E}[R(\boldsymbol{U}) \mid U_j]] = \mathrm{Var}[R_j(U_j)],$$

and

$$V = \mathrm{Var}[R(\boldsymbol{U})] = \sum_{j=1}^{J} \mathrm{Var}[R_j(U_j)] = \sum_{j=1}^{J} V_j.$$

Since we take $R(\boldsymbol{u})$ to be the logarithm of the timing, duration, or strength of the response profile of a molecular species of interest, the requirement that $R(\boldsymbol{u})$ is additive is less restrictive than requiring one of these features to be additive. Indeed, we may have $D(\boldsymbol{u}) = \prod_{j=1}^{J} \exp\{D_j(u_j)\}$ for the duration, in which case $R(\boldsymbol{u}) = \ln D(\boldsymbol{u}) = \sum_{j=1}^{J} D_j(u_j)$ and $R(\boldsymbol{u})$ will be additive.

C. Monte Carlo Estimation

Let X and Y be two statistically independent random variables with probability density functions $f_X(x)$ and $f_Y(y)$, respectively, and $Z = g(X, Y)$ be another random variable, which is a function of X and Y. The k^{th}-order moment $\mathrm{E}\left[Z^k\right]$ of Z is given by

$$\mathrm{E}\left[Z^k\right] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [g(x, y)]^k f_X(x) f_Y(y)\, dxdy.$$

We can derive a Monte Carlo estimator $\widehat{\mathrm{E}}\left[Z^k\right]$ for this quantity by independently drawing samples $\{x^{(l)}, l = 1, 2, \ldots, L\}$ of X and $\{y^{(l)}, l = 1, 2, \ldots, L\}$ of Y from the probability density functions $f_X(x)$ and $f_Y(y)$, respectively, and by setting

$$\widehat{\mathrm{E}}\left[Z^k\right] = \frac{1}{L} \sum_{l=1}^{L} [g(x^{(l)}, y^{(l)})]^k.$$

As a special case, we can estimate the mean and variance of Z by

$$\begin{aligned} \widehat{\mathrm{E}}[Z] &= \frac{1}{L} \sum_{l=1}^{L} g(x^{(l)}, y^{(l)}) \\ \widehat{\mathrm{Var}}[X] &= \frac{1}{L} \sum_{l=1}^{L} g^2(x^{(l)}, y^{(l)}) - \widehat{\mathrm{E}}^2[Z], \end{aligned}$$

since $\mathrm{Var}[Z] = \mathrm{E}\left[Z^2\right] - \mathrm{E}^2[Z]$.

To derive a Monte Carlo estimator for the variance $\mathrm{Var}\big[\mathrm{E}\big[g(X,Y) \mid Y\big]\big]$, note that

$$
\begin{aligned}
&\mathrm{Var}\big[\mathrm{E}\big[g(X,Y) \mid Y\big]\big] \\
&\quad = \int_{-\infty}^{\infty} \mathrm{E}^2\big[g(X,Y) \mid Y\big]\, f_Y(y)\, dy - \big\{\mathrm{E}\big[\mathrm{E}\big[g(X,Y) \mid Y\big]\big]\big\}^2 \\
&\quad = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} g(x_1,y)\, f_X(x_1)\, dx_1\right]\left[\int_{-\infty}^{\infty} g(x_2,y)\, f_X(x_2)\, dx_2\right] f_Y(y)\, dy - \mathrm{E}^2\big[g(X,Y)\big] \\
&\quad = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x_1,y) g(x_2,y)\, f_X(x_1)\, f_X(x_2)\, f_Y(y)\, dx_1 dx_2 dy - \mathrm{E}^2[Z].
\end{aligned}
$$

This leads to the following Monte Carlo estimator:

$$
\widehat{\mathrm{Var}}\big[\mathrm{E}\big[g(X,Y) \mid Y\big]\big] = \frac{1}{L}\sum_{l=1}^{L} g(x_1^{(l)}, y^{(l)}) g(x_2^{(l)}, y^{(l)}) - \widehat{\mathrm{E}}^2[Z],
$$

where $\{x_1^{(l)}, l = 1,2,\ldots,L\}$, $\{x_2^{(l)}, l = 1,2,\ldots,L\}$ are two sets of samples of X drawn independently from the probability density function $f_X(x)$, and $\{y^{(l)}, l = 1,2,\ldots,L\}$ is a set of samples of Y drawn independently from the probability density function $f_Y(y)$.

D. Sketch of Proof for Conditions C.1-C.7

In this section, we show that conditions C.1–C.7 are indeed satisfied by the Monte Carlo variance estimators introduced in Section V of the Main Text.

Condition C.1 is a direct consequence of the fact that

$$
\widehat{\mathrm{Var}}_j[R(\boldsymbol{U})] = \frac{1}{4L}\Big\{\sum_{l=1}^{L}\big[R(\boldsymbol{u}^{(l)}) - R(\boldsymbol{u}^{(L+l)})\big]^2 + \sum_{l=1}^{L}\big[R(\boldsymbol{u}_j^{(l)}) - R(\boldsymbol{u}_{(j)}^{(l)})\big]^2\Big\}, \tag{5}
$$

and

$$
\widehat{\mathrm{Var}}_{jj'}[R(\boldsymbol{U})] = \frac{1}{4L}\Big\{\sum_{l=1}^{L}\big[R(\boldsymbol{u}_j^{(l)}) - R(\boldsymbol{u}_{(j)}^{(l)})\big]^2 + \sum_{l=1}^{L}\big[R(\boldsymbol{u}_{j'}^{(l)}) - R(\boldsymbol{u}_{(j')}^{(l)})\big]^2\Big\}. \tag{6}
$$

Conditions C.2, C.3, C.5, and C.6 can be shown from Eq. (5) and Eq. (6) above, the fact that

$$\widehat{\mathrm{Var}}\big[\mathrm{E}\big[R(\boldsymbol{U}) \mid U_j\big]\big] = \frac{1}{2L}\Big\{\sum_{l=1}^{L}\big[R(\boldsymbol{u}^{(l)}) - R(\boldsymbol{u}^{(l)}_{(j)})\big]\big[R(\boldsymbol{u}^{(l)}_{j}) - R(\boldsymbol{u}^{(L+l)})\big]\Big\}$$

$$\widehat{\mathrm{Var}}\big[\mathrm{E}\big[R(\boldsymbol{U}) \mid \boldsymbol{U}_{(j)}\big]\big] = \frac{1}{2L}\Big\{\sum_{l=1}^{L}\big[R(\boldsymbol{u}^{(l)}) - R(\boldsymbol{u}^{(l)}_{j})\big]\big[R(\boldsymbol{u}^{(l)}_{(j)}) - R(\boldsymbol{u}^{(L+l)})\big]\Big\}$$

$$\widehat{\mathrm{Var}}\big[\mathrm{E}\big[R(\boldsymbol{U}) \mid U_j, U_{j'}\big]\big] = \frac{1}{2L}\Big\{\sum_{l=1}^{L}\big[R(\boldsymbol{u}^{(l)}_{j}) - R(\boldsymbol{u}^{(l)}_{j'})\big]\big[R(\boldsymbol{u}^{(l)}_{(j')}) - R(\boldsymbol{u}^{(l)}_{(j)})\big]\Big\},$$

and the facts that fixing U_j implies that $\boldsymbol{u}^{(l)} = \boldsymbol{u}^{(l)}_{(j)}$ and $\boldsymbol{u}^{(L+l)} = \boldsymbol{u}^{(l)}_{j}$, fixing $\boldsymbol{U}_{(j)}$ implies that $\boldsymbol{u}^{(l)} = \boldsymbol{u}^{(l)}_{j}$ and $\boldsymbol{u}^{(L+l)} = \boldsymbol{u}^{(l)}_{(j)}$, whereas, fixing $\boldsymbol{U}_{(j,j')}$ implies that $\boldsymbol{u}^{(l)}_{j} = \boldsymbol{u}^{(l)}_{(j')}$ and $\boldsymbol{u}^{(l)}_{(j)} = \boldsymbol{u}^{(l)}_{j'}$. To show Condition C.4, we define

$$\begin{aligned}
h_1(l) &:= R(\boldsymbol{u}^{(l)}) \\
h_2(l) &:= R(\boldsymbol{u}^{(L+l)}) \\
h_3(l) &:= R(\boldsymbol{u}^{(l)}_{j}) \\
h_4(l) &:= R(\boldsymbol{u}^{(l)}_{(j)}).
\end{aligned}$$

Then, by using some straightforward algebra, we can show that

$$\{[h_1(l) + h_2(l)] - [h_3(l) + h_4(l)]\}^2 \geq 0$$

implies

$$\begin{aligned}
\frac{1}{4}\Big[h_1^2(l) + h_2^2(l) + h_3^2(l) + h_4^2(l)\Big] - \frac{1}{2}\left[h_1(l)h_2(l) + h_3(l)h_4(l)\right] \\
\geq \frac{1}{2}\left[h_1(l)h_3(l) + h_2(l)h_4(l)\right] - \frac{1}{2}\left[h_1(l)h_2(l) + h_3(l)h_4(l)\right] \\
+ \frac{1}{2}\left[h_1(l)h_4(l) + h_2(l)h_3(l)\right] - \frac{1}{2}\left[h_1(l)h_2(l) + h_3(l)h_4(l)\right],
\end{aligned}$$

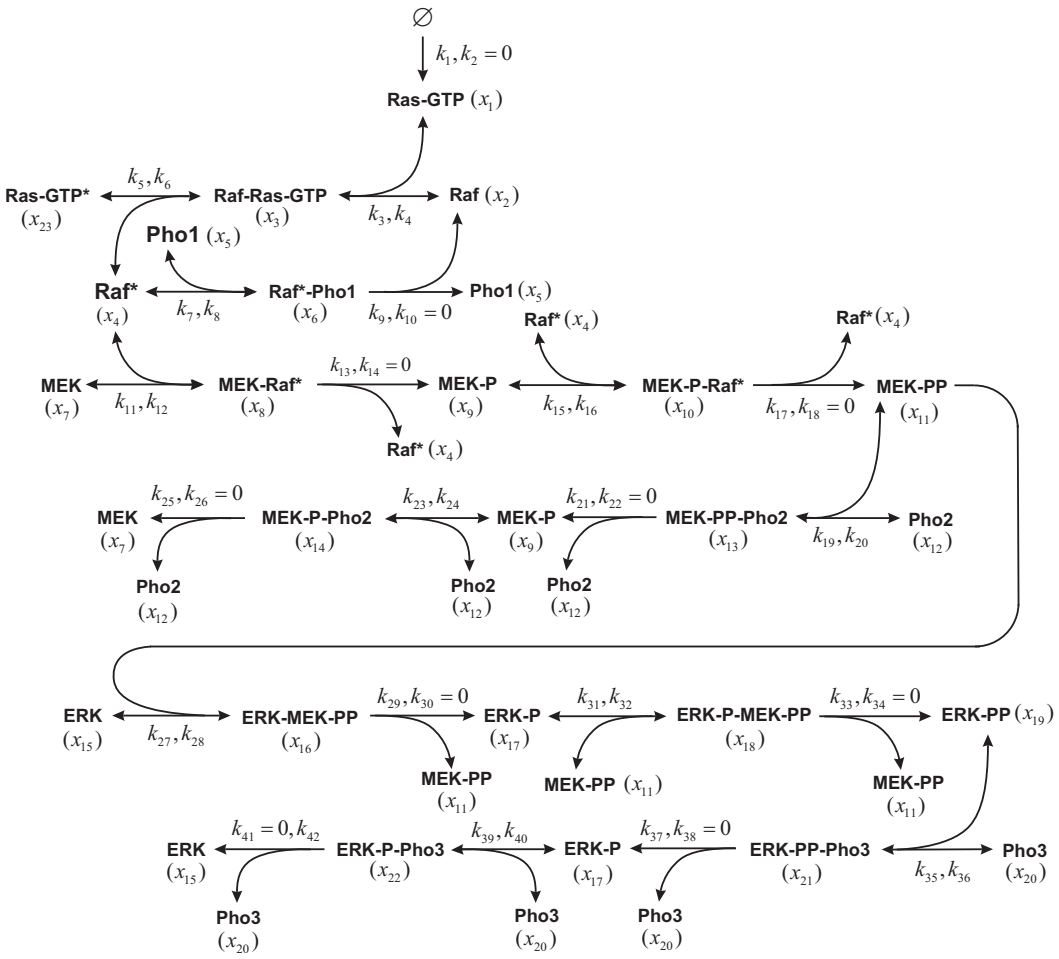

Figure 1: A biochemical reaction model of the MAPK signaling cascade.

which in turn implies that $\widehat{\mathrm{Var}}_j\left[R(\boldsymbol{U})\right] \geq \widehat{\mathrm{Var}}\left[\mathrm{E}\left[R(\boldsymbol{U}) \mid U_j\right]\right] + \widehat{\mathrm{Var}}\left[\mathrm{E}\left[R(\boldsymbol{U}) \mid \boldsymbol{U}_{(j)}\right]\right]$. Condition C.7 can be shown similarly.

E. MAPK Signaling Cascade Model

In this section, we list the biochemical reactions associated with the MAPK signaling cascade model we consider in the Main Text and provide nominal values for the reaction rate constants and values for the initial molecular activities. We depict this model in Fig. 1, whereas, in Fig. 2, we depict the activity profiles of selected species. We have adopted the data from Schoeberl *et al.*,[9] with a few rate constant values updated from the 'JWS Online Cellular Systems Modeling' web

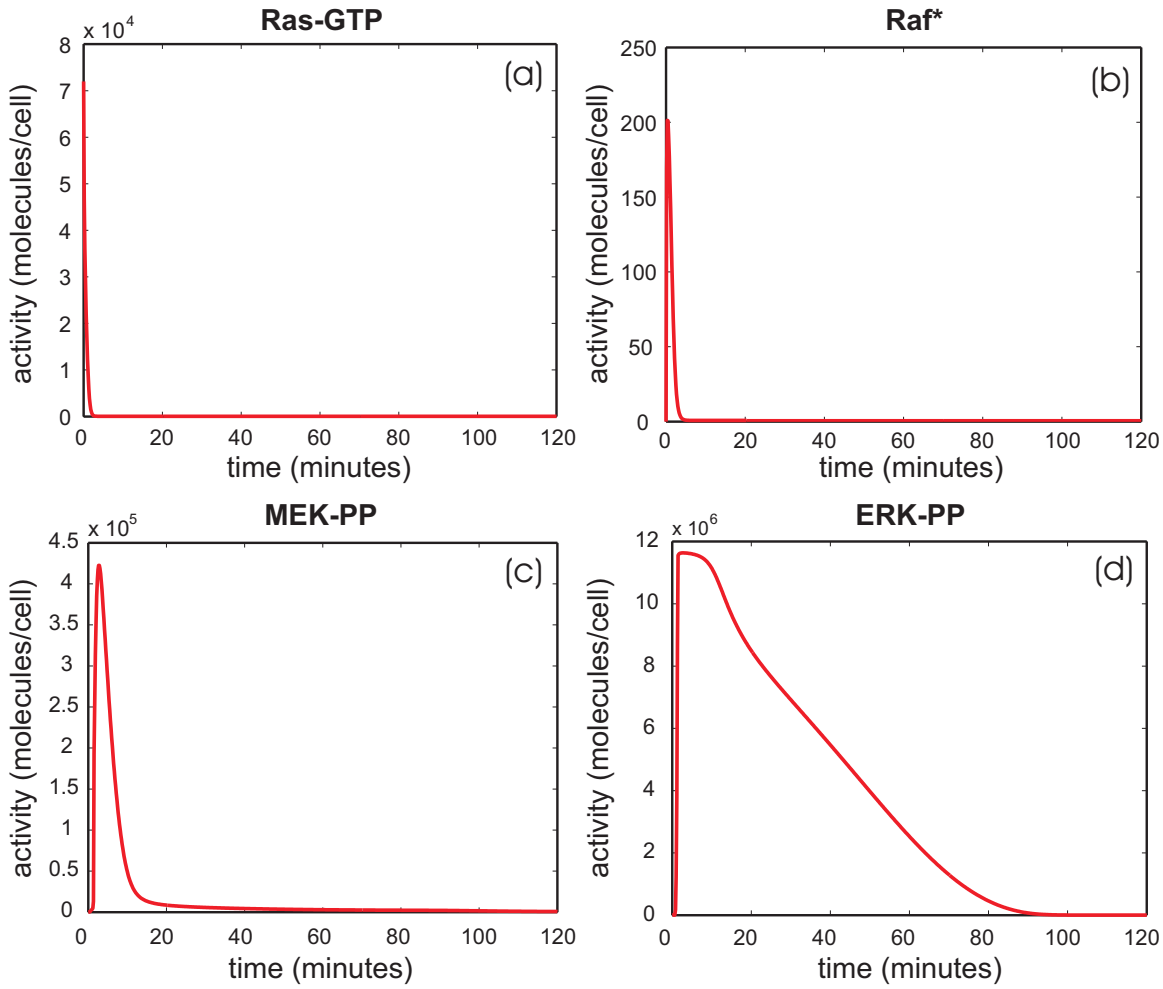

Figure 2: Activity profiles of: (a) Ras-GTP, (b) Raf*, (c) MEK-PP, and (d) ERK-PP, predicted by the MAPK cascade model depicted in Fig. 1.

site (http://jjj.biochem.sun.ac.za). The first reaction in the model depicted in Fig. 1 compensates for Ras-GTP synthesis, which, in reality, is accomplished by a complex epidermal growth factor (EGF)-induced signalling pathway.[9] We have set the reaction rate constant of Ras-GTP synthesis equal to $3s^{-1}$. This value results in an ERK-PP activity profile that is similar to the one reported by Schoeberl *et al.*,[9] with 50ng/ml EGF; compare Fig. 2(d) with Fig. 2(F) in Schoeberl *et al.*[9]

Reactions

No.	Reaction	Rate Constant (s^{-1})
1	$\emptyset \rightarrow$ Ras-GTP	$\kappa_1 = 3$
	Ras-GTP $\rightarrow \emptyset$	$\kappa_2 = 0$
2	Ras-GTP + Raf $\rightarrow$ Raf-Ras-GTP	$\kappa_3 = 1.6605 \times 10^{-6}$
	Raf-Ras-GTP $\rightarrow$ Raf + Ras-GTP	$\kappa_4 = 5.3 \times 10^{-3}$
3	Raf-Ras-GTP $\rightarrow$ Raf* + Ras-GTP*	$\kappa_5 = 1$
	Raf* + Ras-GTP* $\rightarrow$ Raf-Ras-GTP	$\kappa_6 = 1.1624 \times 10^{-6}$
4	Raf* + Pho1 $\rightarrow$ Raf*-Pho1	$\kappa_7 = 1.1790 \times 10^{-4}$
	Raf*-Pho1 $\rightarrow$ Raf* + Pho1	$\kappa_8 = 0.2$
5	Raf*-Pho1 $\rightarrow$ Raf + Pho1	$\kappa_9 = 1$
	Raf + Pho1 $\rightarrow$ Raf*-Pho1	$\kappa_{10} = 0$
6	MEK + Raf* $\rightarrow$ MEK-Raf*	$\kappa_{11} = 1.9428 \times 10^{-5}$
	MEK-Raf* $\rightarrow$ MEK + Raf*	$\kappa_{12} = 3.3 \times 10^{-2}$
7	MEK-Raf* $\rightarrow$ MEK-P + Raf*	$\kappa_{13} = 3.5$
	MEK-P + Raf* $\rightarrow$ MEK-Raf*	$\kappa_{14} = 0$
8	MEK-P + Raf* $\rightarrow$ MEK-P-Raf*	$\kappa_{15} = 1.9428 \times 10^{-5}$
	MEK-P-Raf* $\rightarrow$ MEK-P + Raf*	$\kappa_{16} = 3.3 \times 10^{-2}$
9	MEK-P-Raf* $\rightarrow$ MEK-PP + Raf*	$\kappa_{17} = 2.9$
	MEK-PP + Raf* $\rightarrow$ MEK-P-Raf*	$\kappa_{18} = 0$
10	MEK-PP + Pho2 $\rightarrow$ MEK-PP-Pho2	$\kappa_{19} = 2.3746 \times 10^{-5}$
	MEK-PP-Pho2 $\rightarrow$ MEK-PP + Pho2	$\kappa_{20} = 0.8$
11	MEK-PP-Pho2 $\rightarrow$ MEK-P + Pho2	$\kappa_{21} = 5.8 \times 10^{-2}$
	MEK-P + Pho2 $\rightarrow$ MEK-PP-Pho2	$\kappa_{22} = 0$
12	MEK-P + Pho2 $\rightarrow$ MEK-P-Pho2	$\kappa_{23} = 4.4835 \times 10^{-7}$
	MEK-P-Pho2 $\rightarrow$ MEK-P + Pho2	$\kappa_{24} = 0.5$
13	MEK-P-Pho2 $\rightarrow$ MEK + Pho2	$\kappa_{25} = 5.8 \times 10^{-2}$
	MEK + Pho2 $\rightarrow$ MEK-P-Pho2	$\kappa_{26} = 0$
14	ERK + MEK-PP $\rightarrow$ ERK-MEK-PP	$\kappa_{27} = 8.8673 \times 10^{-5}$
	ERK-MEK-PP $\rightarrow$ ERK + MEK-PP	$\kappa_{28} = 1.833 \times 10^{-2}$
15	ERK-MEK-PP $\rightarrow$ ERK-P + MEK-PP	$\kappa_{29} = 16$
	ERK-P + MEK-PP $\rightarrow$ ERK-MEK-PP	$\kappa_{30} = 0$
16	ERK-P + MEK-PP $\rightarrow$ ERK-P-MEK-PP	$\kappa_{31} = 8.8673 \times 10^{-5}$
	ERK-P-MEK-PP $\rightarrow$ ERK-P + MEK-PP	$\kappa_{32} = 1.833 \times 10^{-2}$
17	ERK-P-MEK-PP $\rightarrow$ ERK-PP + MEK-PP	$\kappa_{33} = 5.7$
	ERK-PP + MEK-PP $\rightarrow$ ERK-P-MEK-PP	$\kappa_{34} = 0$
18	ERK-PP + Pho3 $\rightarrow$ ERK-PP-Pho3	$\kappa_{35} = 2.3414 \times 10^{-5}$
	ERK-PP-Pho3 $\rightarrow$ ERK-PP + Pho3	$\kappa_{36} = 0.6$
19	ERK-PP-Pho3 $\rightarrow$ ERK-P + Pho3	$\kappa_{37} = 0.246$
	ERK-P + Pho3 $\rightarrow$ ERK-PP-Pho3	$\kappa_{38} = 0$
20	ERK-P + Pho3 $\rightarrow$ ERK-P-Pho3	$\kappa_{39} = 8.3027 \times 10^{-6}$
	ERK-P-Pho3 $\rightarrow$ ERK-P + Pho3	$\kappa_{40} = 0.5$
21	ERK + Pho3 $\rightarrow$ ERK-P-Pho3	$\kappa_{41} = 0$
	ERK-P-Pho3 $\rightarrow$ ERK + Pho3	$\kappa_{42} = 0.246$

Initial Conditions

No.	species	molecules/cell
1	Ras-GTP	7.20×10^4
2	Raf	4.00×10^4
3	Raf-Ras-GTP	0
4	Raf*	0
5	Pho1	4.00×10^4
6	Raf*-Pho1	0
7	MEK	2.10×10^8
8	MEK-Raf*	0
9	MEK-P	0
10	MEK-P-Raf*	0
11	MEK-PP	0
12	Pho2	4.00×10^4
13	MEK-PP-Pho2	0
14	MEK-P-Pho2	0
15	ERK	2.21×10^7
16	ERK-MEK-PP	0
17	ERK-P	0
18	ERK-P-MEK-PP	0
19	ERK-PP	0
20	Pho3	1.00×10^7
21	ERK-PP-Pho3	0
22	ERK-P-Pho3	0
23	Ras-GTP*	0

To use the previous reaction rate values, we must check whether these values satisfy Eq. (8) in the Main Text. In particular, we need to show that there exist capacities $c_m^{\ddagger}$, $m = 1, 2, \ldots, M$, and c_n, $n = 1, 2, \ldots, N$, so that Eq. (8) in the Main Text is satisfied. Note that

$$\frac{\kappa_{2m-1}}{\kappa_{2m}} = \prod_{n=1}^{N} c_n^{s_{nm}}, \quad \text{for every } \; m \in \mathcal{M}_r, \tag{7}$$

from Eq. (19) in the Main Text, where s_{nm} is given by Eq. (3) in the Main Text and $\mathcal{M}_r$ denotes the set of all reversible reactions in the MAPK cascade (i.e., reactions 2–4, 6, 8, 10, 12, 14, 16, 18, and 20). If we denote by $\mathbb{S}_r$ the $N \times M_r$ stoichiometry matrix associated with the reversible reactions, by $\boldsymbol{c}$ the $N \times 1$ vector of the molecular capacities $c_n, n = 1, 2, \ldots, N$, and by $\boldsymbol{r}$ the $M_r \times 1$ vector of the reaction rate ratios $\kappa_{2m-1}/\kappa_{2m}$, $m \in \mathcal{M}_r$, then we can write Eq. (7) above in the following matrix-vector form:

$$\mathbb{S}_r^T \ln \boldsymbol{c} = \ln \boldsymbol{r},$$

where $\ln \boldsymbol{u}$ denotes a vector with elements $\ln u_i$, where u_i is the i^{th} element of vector $\boldsymbol{u}$.

It turns out that, for the MAPK cascade, the columns of the stoichiometry matrix $\mathbb{S}_r$ are linearly independent; i.e., $\text{rank}(\mathbb{S}_r) = M_r$. As a consequence, we can write

$$\mathbb{S}_r^T \ln \boldsymbol{c} = \begin{bmatrix} \mathbb{S}_* \\ \mathbb{S}_{**} \end{bmatrix}^T \begin{bmatrix} \ln \boldsymbol{c}_1 \\ \ln \boldsymbol{c}_2 \end{bmatrix} = \mathbb{S}_*^T \ln \boldsymbol{c}_1 + \mathbb{S}_{**}^T \ln \boldsymbol{c}_2 = \ln \boldsymbol{r}, \tag{8}$$

where $\mathbb{S}_*$ is the $M_r \times M_r$ matrix that contains all linearly independent rows of $\mathbb{S}_r$, $\mathbb{S}_{**}$ is the $(N - M_r) \times M_r$ matrix that contains the remaining rows of $\mathbb{S}_r$, and $\boldsymbol{c}_1$, $\boldsymbol{c}_2$ are the corresponding subvectors of $\boldsymbol{c}$. Note that $\mathbb{S}_*$ is an invertible matrix. Therefore, Eq. (8) above implies that

$$\ln \boldsymbol{c}_1 = (\mathbb{S}_*^T)^{-1} \left(\ln \boldsymbol{r} - \mathbb{S}_{**}^T \ln \boldsymbol{c}_2 \right).$$

We can now set $\ln \boldsymbol{c}_2 = 0$, in which case $\ln \boldsymbol{c}_1 = (\mathbb{S}_*^T)^{-1} \ln \boldsymbol{r}$. Therefore, given the reaction rate constants of the reversible reactions, we can find capacity values for all molecular species in the system. Subsequently, we can determine the capacities of the activated complexes by setting

$$c_m^{\ddagger} = \kappa_{2m-1} \frac{h}{k_B T} \prod_{n=1}^{N} c_n^{\nu_{nm}}, \quad \text{for } m = 1, 2, \ldots, M.$$

The previous discussion shows that the rate constant values we are using in this paper correspond to a thermodynamically feasible model for the MAPK cascade, since, given these values, we can find appropriate capacity values so that the Eyring-Polanyi equations are satisfied. Although we could use the previous steps to calculate the actual capacity values, we do not do that here since these values are immaterial for sensitivity analysis.

References

(1) Marshall, C. J. *Cell* **1995**, *80*, 179–185.

(2) Murphy, L. O.; Smith, S.; Chen, R.-H.; Fingar, D. C.; Blenis, J. *Nature Cell Biology* **2002**, *4*, 556–564.

(3) Murphy, L. O.; MacKeigan, J. P.; Blenis, J. *Molecular and Cellular Biology* **2004**, *24*, 144–153.

(4) Mayawala, K.; Gelmi, C. A.; Edwards, J. S. *Biophysical Journal: Biophysical Letters* **2004**, L01–L02.

(5) Ebisuya, M.; Kondoh, K.; Nishida, E. *Journal of Cell Science* **2005**, *118*, 2997–3002.

(6) Tombes, R. M.; Auer, K. L.; Mikkelsen, R.; Valerie, K.; Wymann, M. P.; Marshall, C. J.; McMahon, M.; Dent, P. *The Biochemical Journal* **1998**, *330*, 1451–1460.

(7) Asthagiri, A. R.; Reinhart, C. A.; Horwitz, A. F.; Lauffenburger, D. A. *Journal of Cell Science* **2000**, *113*, 4499–4510.

(8) Traverse, S.; Seedorf, K.; Paterson, H.; Marshall, C. J.; Cohen, P.; Ullrich, A. *Current Biology* **1994**, *4*, 694–701.

(9) Schoeberl, B.; Eichler-Jonsson, C.; Gilles, E. D.; Müller, G. *Nature Biotechnology* **2002**, *20*, 370–375.